Estimação Intervalar: A Lógica da Quantidade Pivotal

ESTAT0078 – Inferência I

Prof. Dr. Sadraque E. F. Lucena

sadraquelucena@academico.ufs.br

http://sadraquelucena.github.io/inferencia1

UNIVERSIDADE FEDERAL DE SERGIPE

Introdução

- **Estimativa Pontual ($\hat{\theta}$):**
  - Um único valor numérico baseado na amostra.
  - *Limitação:* Não indica precisão ou incerteza da estimativa.
- **Estimativa Intervalar:**
  - Um intervalo de valores plausíveis $[L(\mathbf{X}), U(\mathbf{X})]$.
  - Os limites são funções da amostra (variáveis aleatórias).
- **Exemplo:** Se você estima que o custo médio de um sinistro de uma seguradora é R$ 1.000,00 e baseia todo o seu fundo de reserva nisso, qualquer variação para cima quebra a empresa.
- O **Intervalo de Confiança** quantifica essa incerteza, permitindo uma gestão de riscos mais robusta.

Objetivo da Estimação Intervalar

- **Construção do Intervalo:**
  - Encontrar duas estatísticas, $L(\mathbf{X})$ e $U(\mathbf{X})$, tais que:

$$P[L(\mathbf{X}) \leq \theta \leq U(\mathbf{X})] = 1 - \alpha$$

- **Nível de Confiança $(1 - \alpha)$:**
  - É a probabilidada priori (antes da coleta dos dados) de que o intervalo contenha o verdadeiro valor do parâmetro θ.
- **Significância (α):**
  - A probabilidade de erro, ou seja, do intervalo construído não conter o parâmetro.

Definição 16.1: Quantidade Pivotal

Uma variável aleatória $Q(X_1, \ldots, X_n; \theta) = Q(\mathbf{X}; \theta)$ é dita ser uma quantidade pivotal para o parâmetro θ se sua distribuição de probabilidade **não depende** de θ.

- **Diferença Fundamental:**
  - **Estatística:** Não pode conter θ. Conseguimos calcular o seu valor numérico apenas com os dados da amostra. (Ex: $\overline{X}$).
  - **Quantidade Pivotal:** Deve conter o parâmetro, mas sua distribuição deve ser livre dele (Ex: $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$).

Exemplo 16.1

Considere $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$. As equações abaixo são quantidades pivotais para o parâmetro de interesse?

a. $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. (Dica: A distribuição depende de μ?)

b. $S^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$. (Dica: A distribuição depende de σ^2?)

c. $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ (com σ conhecido).

d. $Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.

e. $D = \bar{X} - \mu \sim N(0, \sigma^2/n)$

Passos para o Intervalo de Confiança

- **Passo 1:** Selecione uma **Quantidade Pivotal** $Q(\mathbf{X}, \theta)$ com distribuição conhecida.
- **Passo 2:** Escolha o nível de confiança $1 - \alpha$ (ex: 95%) e encontre os valores críticos λ_1 e λ_2 na tabela, tais que:

$$P(\lambda_1 \leq Q(\mathbf{X}, \theta) \leq \lambda_2) = 1 - \alpha.$$

- **Passo 3:** Use álgebra para isolar o parâmetro θ no centro da desigualdade.
- **Resultado:** $P(L(\mathbf{X}) \leq \theta \leq U(\mathbf{X})) = 1 - \alpha.$
  - $L(\mathbf{X})$ e $U(\mathbf{X})$ são os limites do intervalo de confiança.

Exemplo 16.2

Seja $X_1, \ldots, X_n$ uma a.a. de $X \sim N(\mu, \sigma^2)$ com σ^2 conhecido. Encontre o IC para μ usando a quantidade pivotal $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

Exemplo 16.3

Sejam $X_1, \ldots, X_n \sim Exp(\theta)$. Utilize a quantidade pivotal $Q = 2\theta \sum X_i \sim \chi^2_{2n}$ para derivar o IC para θ.

Exemplo 16.4

Seja $X \sim U(0, \theta)$ uma única observação. Use a quantidade pivotal $Q = X/\theta \sim U(0, 1)$ e monte um IC de 95% para θ.

A Armadilha da Interpretação

O nível de confiança $(1 - \alpha)$ mostra em que porcentagem das vezes o seu cálculo vai produzir um resultado correto se você repetir o experimento várias vezes.

1. **O Intervalo é Aleatorio**

- **Antes da coleta:** Os limites $L(\mathbf{X})$ e $U(\mathbf{X})$ são funções da amostra (variáveis aleatórias).
- **Depois da coleta:** O intervalo torna-se fixo (ex: $[10, 20]$). Aqui, a probabilidade de conter θ é 0 ou 1 (ou está lá, ou não está).

A Armadilha da Interpretação

2. **O Parâmetro θ é Fixo**

- θ é uma constante desconhecida. Ele não “flutua”. É o intervalo que “tenta cercar” o parâmetro.

A Armadilha da Interpretação

3. **Certo vs. Errado (Exemplo: Confiança de 95%)**

Interpretação	Frase
ERRADO	"Existe 95% de probabilidade de θ estar entre 10 e 20."
CORRETO	"Este intervalo foi construído por um método que, em 95% das vezes, produz intervalos que contêm o verdadeiro parâmetro."
CORRETO	"Temos 95% de confiança de que o parâmetro θ esteja no intervalo [10,20]."

Analogia do Alvo e o Aro

Imagine que o parâmetro θ é um alvo fixo no chão e o Intervalo de Confiança é um aro que você lança.

1. **Antes de lançar (Passo a priori):**

- O aro ainda está na sua mão. A posição onde ele vai cair é uma variável aleatória.
- **Probabilidade:** Existe 95% de chance do aro cair "cercando" o alvo.

Analogia do Alvo e o Aro

2. **Depois que o aro caiu (Passo a posteriori):**

- O aro já está no chão (os dados já foram coletados e o intervalo calculado, ex: $[10, 20]$).
- **Fato:** Ou o aro cercou o alvo, ou não cercou. Não há mais "95% de chance".
- **Confiança:** O "95%" agora refere-se à precisão do seu braço (o método), não à posição do alvo.
- **Conclusão:** O alvo (θ) nunca se mexe. É o aro (o intervalo) que é lançado e pode ou não capturá-lo. Por isso, após os cálculos, trocamos a palavra **Probabilidade** pela palavra **Confiança**.

Fim